GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:43:16 ; Search time 299.73 Seconds
                                          (without alignments)
                                          260.436 Million cell updates/sec

Title: US-09-331-631A-1
Perfect score: 3542
Sequence: 1 MAINTSNLCSLLFLLSLFLL..........SPRSTKQQQPLVSILDFVGF 666

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_15:*
               1: sp_archea:*
               2: sp_bacteria:*
               3: sp_fungi:*
               4: sp_human:*
               5: sp_invertebrate:*
               6: sp_mammal:*
               7: sp_mhc:*
               8: sp_organelle:*
               9: sp_phage:*
               10: sp_plant:*
               11: sp_rodent:*
               12: sp_virus:*
               13: sp_vertebrate:*
               14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	3542	100.0	666	10	Q9SPL5	macadamia i
2	3412	96.3	666	10	Q9SPL4	macadamia i
3	3227	91.1	625	10	Q9SPL3	macadamia i
4	1314.5	37.1	593	10	Q9SEW4	juglans reg
5	1198.5	33.8	810	10	Q9ZW13	cucurbita m
6	1025	28.9	525	10	Q43358	theobroma c
7	898	25.4	582	10	O03865	zea mays (m
8	864.5	24.4	637	10	O03678	hordeum vul
9	842	23.8	613	10	Q9M3X6	pisum sativ
10	812	22.9	544	10	O22120	glycine max
11	802.5	22.7	489	10	Q9SP11	glycine max
12	787	22.2	545	10	O41674	vicia narbo
13	769.5	21.7	483	10	O49927	pisum sativ
14	752	21.2	448	10	O40873	picea glauc
15	749	21.1	486	10	Q9LUJ7	arabidopsis
16	747	21.1	417	10	O22121	glycine max
17	741.5	20.9	438	10	O43626	picea glauc
18	741	20.9	450	10	O40844	picea glauc
19	731.5	20.7	463	10	Q41677	vicia narbo

RESULT   1
Q9SPL5

ID	Q9SPL5	PRELIMINARY;	PRT;	666 AA.

AC   Q9SPL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL;
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7S
RT   globulin protein in Macadamia integrifolia kernels.";
RL   Plant J. 0:0-0(1999).
RL   EMBL; AF161883; AAD54244.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR00113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE   666 AA; 78217 MW; C752B884B2DF0224 CRC64;

ALIGNMENTS

Query Match   100.0%;   Score 3542; DB 10; Length 666;
Best Local Similarity   100.0%;
Matches 666; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MAINTSNLCSLLFLLSLFLLSTTVSLAESFDRQEYEECKRQCMQLETSGQMRKCVSQCD 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAINTSNLCSLLFLLSLFLLSTTVSLAESFDRQEYEECKRQCMQLETSGQMRKCVSQCD 60

Qy   61 KRFEEDIDWSKYDNQEDPQTECQCQCRRCRQQESGPRQQQYCQRRCKEICEEEEYNRQR 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 KRFEEDIDWSKYDNQEDPQTECQCQCRRCRQQESGPRQQQYCQRRCKEICEEEEYNRQR 120

Qy   121 DPQQQYEQCQCKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQREDEEKYEERM 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 DPQQQYEQCQCKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQREDEEKYEERM 180

20	705.5	19.9	518	10	Q9M3X8	lens culina
21	639.5	18.1	461	10	Q9ZRG9	oryza sativ
22	551.5	15.6	232	10	Q9SQ49	rulingia ma
23	549	15.5	430	10	O41115	phaseolus v
24	547.5	15.5	414	10	O41727	zamia furfu
25	545	15.4	430	10	Q43633	phaseolus v
26	539	15.2	421	10	Q43632	phaseolus v
27	537	15.2	239	10	Q9SQ50	abroma augu
28	534.5	15.1	236	10	Q9SQ41	theobroma c
29	533.5	15.1	236	10	Q9SQ47	theobroma g
30	532.5	15.0	236	10	Q9SQ36	theobroma g
31	532.5	15.0	236	10	Q9SQ34	theobroma m
32	531.5	15.0	236	10	Q9SQ40	theobroma c
33	531.5	15.0	236	10	Q9SQ37	theobroma g
34	528.5	14.9	236	10	Q9SQ48	theobroma m
35	527.5	14.9	540	10	O03866	zea mays (m
36	534	14.9	239	10	Q9SQ35	theobroma c
37	527	14.9	239	10	Q9S7V9	theobroma b
38	525	14.8	239	10	Q9SQ42	theobroma c
39	524	14.8	238	10	Q9SQ32	theobroma s
40	523.5	14.8	236	10	Q9SQ33	theobroma m
41	523	14.8	239	10	Q9SQ44	theobroma g
42	521	14.7	235	10	Q9SQ38	theobroma a
43	519.5	14.7	236	10	Q9SQ46	theobroma g
44	518	14.6	239	10	Q9SQ43	theobroma b
45	516	14.6	423	10	Q43617	phaseolus l

```
Qy    1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEECKRQCMQLETSGQMRRCVSQCD   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEECKRQCMQLETSGQMRRCVSQCD   60
```

Query Match   96.38%; Score 3412; DB 10; Length 666;
Best Local Similarity   96.18%; Pred. No. 1.3e-247;
Matches 640; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

```
RESULT 2

ID   Q9SPL4       PRELIMINARY;     PRT;   666 AA.
AC   Q9SPL4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL.
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7s
RT   globulin protein in Macadamia integrifolia.";
RL   Plant J. 0:0-0(1999)
DR   EMBL; AF161884; AAD54245.1; -.
DR   HSSP; P02853; 2PHL.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   INTERPRO; IPR001113; -.
SQ   SEQUENCE   666 AA;  78243 MW;  0ECA22F8710F8A7B CRC64;
```

```
Qy  181 KEEDNKRDPQQREYEDCRRRCEQOEPRQOHQCQLRCREQOQHGIRGGDMMNPQRGGSGRY  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 KEEDNKRDPQQREYEDCRRRCEQOEPRQOHQCQLRCREQOQHGIRGGDMMNPQRGGSGRY  240

Qy  241 EEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA  300

Qy  301 FVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNERL  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 FVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNERL  360

Qy  361 HIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQREGV  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 HIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQREGV  420

Qy  421 IIRASQEQIRELTRDDSESRHHHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR   480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 IIRASQEQIRELTRDDSESRHHHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR   480

Qy  481 QLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 QLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR  540

Qy  541 HEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 HEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR  600

Qy  601 ERNVLQQIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQQQSPRSTKQQQPLVSI  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ERNVLQQIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQQQSPRSTKQQQPLVSI  660

Qy  661 LDFVGF 666
        ||||||
Db  661 LDFVGF 666
```

```
RESULT 3

ID   Q9SPL3       PRELIMINARY;     PRT;   625 AA.
AC   Q9SPL3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR (FRAGMENT).
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL.
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7s
RT   globulin protein in Macadamia integrifolia.";
RL   Plant J. 0:0-0(1999).
DR   EMBL; AF161885; AAD54246.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
FT   NON_TER      1      1.
SQ   SEQUENCE   625 AA;  73586 MW;  415808A89D370296 CRC64;
```

```
Qy   61 KRFEEDIDWSKYDNQEDPQTECQQCQRCRQQOESGPRQOQYCQRRCKEICEEEEEYNRQR  120
        |||||||||||||| ||||:||::||||||||||||||||||||||||||||||||||||
Db   61 KRFEEDIDWSKYDNQDDPQTDCQQCQRCRQQOESGPRQOQYCQRRCKEICEEEEEYNRQR  120

Qy  121 DPQQOYEQCCQKHCQRRETEPRHMQTCQQCQCERRYEKEKRRQQKRYEEQOQREDEEKYEERM  180
        ||||||||||||||||:|:||  || :||||||:||||:|||:||||||||:|||||||
Db  121 DPQQOYEQCCQRCRHETEPRHMQTCQQCQCERRYEKEKRRQQKRYEEQOQREDEEKYEERM  180

Qy  181 KEEDNKRDPQQREYEDCRRRCEQOEPRQOHQCQLRCREQOQHGIRGGDLINPQRGGSGRY  240
        ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db  181 KEEDNKRDPQQREYEDCRRRCEQOEPRQOHQCQLRCREQOQHGRGGDLINPQRGGSGRY  240

Qy  241 EEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA  300
        |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA  300

Qy  301 FVLPTHLDADAILLVTGGRGALKMIHRDRNRESYNLECGDVIRIPAGTTFYLINRDNNERL  360
        |||||||||||||| :||||||||||:|::|||||||||||||||||||||||||||||
Db  301 FVLPTHLDADAILLVTGGRGALKMIHRDRNRESYNLECGDVIRIPAGTTFYLINRDNNERL  360

Qy  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQAERLRGVLGQQOREGV  420
        |||||||||||||||||||||||||:|||||||||||||||||||:|||||:|||||||
Db  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQAERLRGVLGQQOREGV  420

Qy  421 IISASQEQIRELTRDDSERRWHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR   480
        || ||||||||||||||| :||:|||||||||||||||||||||||||||||||||||
Db  421 IISASQEQIRELTRDDSERRWHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR   480

Qy  481 QLQDMDVSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR  540
        |||||| :||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 QLQDMDVSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR  540

Qy  541 HEEEEDVHYEQVGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR              600
        |||||||||||| :|| ::|||||||||||||||||||||||||||
Db  541 HEEEEDVHYEQVRARLSKREAIVVPVGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR  600

Qy  601 ERNVLQQIEPQAMELAFAAPRKEVEELFNSQDESIFFPGPRQHQQQSSRSTKQQQPLVSI  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ERNVLQQIEPQAMELAFAAPRKEVEELFNSQDESIFFPGPRQHQQQSSRSTKQQQPLVSI  660

Qy  661 LDFVGF 666
        ||||||
Db  661 LDFVGF 666
```

Query Match   Best Local Similarity   96.6%;   Score 3227;   DB 10;   Length 625;
Matches   604;   Conservative   9;   Mismatches   12;   Pred. No. 8.8e-234;
Indels   0;   Gaps   0;

```
Qy     42  QCMQLETSGQMRRCVSQCDKRFEEDIDWSKYDNQEDPQTECQQCQORRCRQQESGPRQQQY 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  QCMQLETSGQMRRCVSQCDKRFEEDIDWSKYDNQEDPQTECQQCQORRCRQQESGPRQQQY  60

Qy    102  CQRRCKEICEEEEEYNRQRDPQOQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQ  161
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CQRRCKEICEEEEEYNRQRDPQOQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQ  120

Qy    162  QKRYEEQQQREDEEKYEERMKEEDNKRDPQOQYEQCRRRCEQQREPERQOCOLRCREQOR  221
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  QKRYEEQQQREDEEKYEERMKEGDNKRDPQOQYEDCRRHCEQQEPRLQYQCQRRCQEQOR  180

Qy    222  QHGRGGDMMNPQRGSGRYEEGGEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSK  281
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  QHGRGGDLMNPQRGSGRYEEGEGEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSK  240

Qy    282  LLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVI  341
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVI  300

Qy    342  RIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILAAA  401
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  RIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAA  360

Qy    402  LNTQTEKLRGVFGQQOREGVIIRASQEOIRELTRDDSESRHWHIRRGGESSRGPYNLFNKR 461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  LNTQTERLRGVLQOQOREGVIIRASQEOIRELTRDDSESRRWHIRRGGESSRGPYNLFNKR 420

Qy    462  PLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVASGEADV  521
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  PLYSNKYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASGEADV  480

Qy    522  EMACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVLAGHPVFVSSGNENLL  581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  EMACPHLSGRHGGRGGGKRHEEEEVHYEQVRARLSKREAIVVLAGHPVFVSSGNENLL  540

Qy    582  LFAFGINAQNNHENFLAGRERNVLQOIEPQAMELAFAAAPRKEVEESFNSQODQSTFPPGR 641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  LFAFGINAQONNHENFLAGRERNVLQOIEPQAMELAFAAASRKEVEELFNSQDESIFFPGR 600

Qy    642  QHOQOOSPRSTKQOOPLVSILDFVGF 666
           ||||||||||||||||||||||||||
Db    601  QHOQOOSPRSTKQOOPLVSILDFVGF 625
```

RESULT 4

Q9SEW4

ID   Q9SEW4   PRELIMINARY;   PRT;   593 AA.
AC   Q9SEW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS   Juglans regia (English walnut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fagales; Juglandaceae; Juglans.
OX   NCBI_TaxID=51240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RA   Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT   "Identification and cloning of a cDNA encoding a vicilin-like protein,
RT   Jug r 2, from English walnut kernel (Juglans regia): a major food
RT   allergen.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF066055; AAF18269.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
NON_TER   1   1
FT   1   1
SQ   SEQUENCE   593 AA;   69990 MW;   9BA127E19B18C0D8 CRC64;

Query Match   Best Local Similarity   37.1%;   Score 1314.5;   DB 10;   Length 593;
Matches   271;   Conservative   121;   Mismatches   169;   Pred. No. 2e-90;
Indels   57;   Gaps   16;

```
Qy     73  DNQEDPQTECQQCQORRCRQQESGPRQQQYCQRRCKEICEE----EEEYNRQ-------- 119
           :||| ||:::::|||| ||::|| |||::|||||||||||||    :||||||
Db      9  ENPRDPREQYRQCQEYCRROGOGOQRQOQOOQQCQIRCEERLEDOQRSQEQERERRGRDVDDQN  68

Qy    120  -RDPQOQYEQCQOKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQQREDEEKYEE 178
           :||:|:| ||:::||::| ||| |||: || ||| : :|: :|: ||| ::::
Db     69  PRDPEQRYEQCCQOQOCE-RQRQGEQTLCRRCEQQRQOQEEREQRQRDRQ-------- 117

Qy    179  RMKEEDNKRDPQOQYEVEDCRRRC--EQOEPRQOHCQOLRC-REQORQHGR-GGDMMNPQOR 234
           :|:|:|| |||| ||:::|::| |  :||:||| :|::| | ||:||||| :||:|||
Db    118  -----DPQO-QYHRCQORRCQIOEQOSPEROROCQORQOEREROQRGRDQ 167

Qy    235  GGSGRYEEGGEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLL  294
           :|:|||||::|:|:|||||||||||||||||||:|||||||||||||:||| ||||||||
Db    168  ESRGR---EEQORHNPYYFHSQSIRSRHESEEGEVKYLERRTERTELLRGTENYRVVIL  224

Qy    295  EANPNAFVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINR  354
           |||||||||||||:| |||||||:|||||||::||||:|::|||||||||||||||||:|
Db    225  DANPDNTSMLPHHKDAESVAVVTRGRATLTLVSQETRESFNLECGDVIRVPAGATVVVINQ  284

Qy    355  DNNERLHIAKFLQTISTPGQYKEFFPAGGQNP-EPYLSTFSKEILEAALNTQTEKLRGVF  413
           :|:|:| ||||||:|||||:||:||||||||  |||||||:||:|::|: |:::|||||
Db    285  DSNERLEMVKLLQPVNNPGGQPREYYAAGAKSPDQSYLRVESNDILVAALNTPRDRLERFF  344

Qy    414  GQ--QREGVIIRASQEOIRELTRD--DSESRHWHIRRGGESSRGPYNLFNKRPLYSNKYG  469
           |:  :|||||||||:|| |::|:|  ||:||::||||||::|| ||:|| ::|||||||
Db    345  DQOEQOREGVIIRASQEKLRALSQHAMSAGOQRPW----GRRSSGGPISLKSESPSYSNQFG  400

Qy    470  QAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVASGEADVEMACPHLS  529
           |||:|||||||:|||:|:||:|||:|||||||||||||::|| :||:|||::| | |||
Db    401  QFFEACPEEHRQLQEMDVLVNYAEIKRGAMMVPHINSKATVVVVVEGTGRYEMACPHVS  460

Qy    530  GR-HGGRGGGKRHEEEEDVHYEQVRARLSKREAIVLAGHPVFVSSGNENLLLFAFGIN  588
           || |||:||||||||||:||:||||| |::| ||:|||||::||||:||||||||||:
Db    461  SOSYEGQGRREQOEEESTGRFQKVTARLARGDIFVTPAGHPIAITASQONENLRLLGFDIN  520

Qy    589  AQNNHENFLAGRERNVLQOIEPQAMELAFAAAPRKEVEESNSQODQSIFFPGPRQHOQOQSP 648
           ||:|||||||||||||||:||||:|||||||::|::|||:|||::||:|| ||::||::|:
Db    521  GENNQRDFLAG--QNNINQLEREAKELSFNMPREELEEIFESQMESYFVPPER----QSR  575

Qy    649  RSTKQOOPLVSILDFVGF 666
           ::|| |||||||||:|:|
Db    576  RGQGRDHPLASILDFAFF 593
```

RESULT 5

Q92WI3

ID   Q92WI3   PRELIMINARY;   PRT;   810 AA.
AC   Q92WI3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
PV100.
OS   Cucurbita maxima (Pumpkin) (Winter squash).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Cucurbitales; Cucurbitaceae; Cucurbita.
OX   NCBI_TaxID=3661;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KUROKAWA AMAKURI NANKIN; TISSUE=COTYLEDON;
RX   MEDLINE=99107919; PubMed=9891029;
RA   Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;
RT   "Multiple functional proteins are produced by cleaving Asn-Gln bonds

```
RT     of a single precursor by vacuolar processing enzyme.";
RL     J. Biol. Chem. 274:2563-2570(1999).
DR     EMBL; AB019195; BAA34056.1; -.
HSSP;  P02853; 2PHL.
DR     INTERPRO; IPR001113; -.
PFAM;  PF00546; Seedstore_7s; 1.
PRODOM; PD081059; -; 1.
DR
SQ     SEQUENCE   810 AA;  97314 MW;  A829A3F5742266AB CRC64;

Query Match             33.8%;
Best Local Similarity   34.5%;
Matches  286;  Conservative  141;

Qy      1  MAINTSNICSLLFLLSLFLLSTTVSLAE-------------------------------SEF-  31
           |:::                     ||:|                                  ||
Db      1  MALSKVKLRLCLLAFTLFLACLSVGLGDKGESLSSGAGVDHDGCVNRCEELKGKNVDEFA  60

Qy     32  ----------DRQEYEECKRQCMQLETS-GQMRRCVSQCQDKRFEE-------------  65
                      ||||:|||:|||:    :||:|:|:|:|:||||
Db     61  ACKKGCGVNQRGSPRAEYEVCRLRCQVAERGVEQRQRCEQVCEERLREREQGRGEDVDEV  120

Qy     66  --DIDWSKY----------------------DNQEDPQTECQQCCQRCQQE  93
             :|:|                         |:||:|| |||| |||| |:
Db    121  ERRDPEWEREEQQRRREHEEREHEREERRRERERERERGRGRRDRENERDPKREQEEERQR--REQE  178

Qy     94  SGPRQQQYCQRRCKEICEEEEEYNRQRDPQQQYEQCQKHCQRRETEPRHMQTCQQRCERR  153
           |:|:    |  :|:||:|||||||::|| :||||| ||:||||:||:||::|:||| :||
Db    179  RRRREQEQRERERGERDEEDDEN-QRDPDWRREQ---------ERREQERRREQCERRERQ  232

Qy    154  Y--------EKE-----------KRKQQKRYEQQQRREEDEEKYEERMKEEDNKRDP---  195
                     |:|           :||  |:|| |||:::    :|:||||  |
Db    233  RRGGRGRDDEDENQRDPDWRREQKRREQCEQERRERGRGRDDEDENQRDPDWRREQERREQ  292

Qy    196  DCRRRCEQQE------------------PRQQHQQCQLRCREQQRQHGRGRG---  227
           |:||    |:                  ||  | :|::|||:|:|
Db    293  ERRRRREREQEQRERERREEHRGRGRDVEDENQRDPDWRREQERRERERERERGRGRDGRGRE  352

Qy    228  ----------DMMNPQRGGSGRYEE----------GEEEQSDNPYYFPERSLSTRFRTEEGHISVL  273
                      |:::||:| |||             |::||||:||:|||:|::|||:|||||:|||
Db    353  EQQRSREDDERRREROHGGRSRVNQVAIRRTEQQEQQSNNYYFQRQRFQSRYRSDEGHWRVL  412

Qy    274  ENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADADILLVIGGRGALKMIHHDNR---  330
           || |||  |: ||||||:|||||||||||||||||||  | |:|:||| |::|:|:|
Db    413  ERFSERSELLKGIKNQRLALLEARPHFFIVPHHLDAECVLIVVRGRATITTVVQEKRETR  472

Qy    331  -ESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPY  389
            |:::| ||| :| ||||||:|:|| |:|||||||||:|||| | |||||||||:|:|||
Db    473  KESYNVESGDVMTIPAGTTLYLANQE-NEDLQIVKLVQPVNNPGEFKDYLSRAGGES-QAY  530

Qy    390  LSTFSKEILEAALNTQTEKLRGVFGQQRE--GVIIRASQEQIRELTRDDSESRHWHIRRG  447
           |||||||||||||:||||||||||||||:   |||||||||||:|| ||||||||:|:||
Db    531  YSVFSNDVLEAAALNIPRDKLERIFKQRRERGKIVRASQEQLRALSQ----RATSVRKG  585

Qy    448  GESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTR  507
           |||||||||:|||||||||||||||||| ||::|||||||||||||||||||||||||||
Db    586  SRGVRAPIKLESQTPVYNNQYGQMFEACPDEEPQLRRTDVAISVVDLKQGGMMVPHFNSR  645

Qy    508  STKVVVVASGEADVEMACPHLSGRHGGRGGGK-----RHEEED-----VHYEQVRARLS  557
           |||||||:|||:|:||||||||||:||||:||      ||:|::    |||:|||||||
Db    646  ATWVFVFVSEGAGSFEMACPHIQSSWQRGRREEEHWRREEEERREERSGRFERVAGRLS  705

Qy    558  KREAIVVLAGHPVFVVSSGGNENLLFAFGINAQNNHENFLAGRERRNVLQQIEPQAMELAF  617
           ||||:|||||||||||||||||:||:||||||||||:|||| |::|:|  || | ||:|||
Db    706  EGGVLVIPAGHPIAIAMASPNENLRLVGFGINAENNHRNFLAGRE-NIMNELDREAKELAF  764

Qy    618  AAPRKEVEESNSQDQSIFFPGPRQHQQQSPRSTKQQQPLVSILDFVGF  666
           |:||:  ||||||||:|||||||  |||||||||| ||||||||||||:
Db    765  NVEGKQADEIFRSQRESFFTEGP----EGGRRRSTERSPLLSILKLAGY  809
```

```
ID      Q43358              PRELIMINARY;       PRT;      525 AA.
AC      Q43358;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE      VICILIN PRECURSOR.
GN      CSV.
OS      Theobroma cacao (Cacao).
OC      Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC      Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC      Malvales; Malvaceae; Theobroma.
OX      NCBI_TaxID=3641;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=LEAVES;
RX      MEDLINE=92288309; PubMed=1600151;
RA      McHenry L., Fritz P.J.;
RT      "Comparison of the structure and nucleotide sequences of vicilin genes
RT      of cocoa and cotton raise questions about vicilin evolution.";
RL      Plant Mol. Biol. 18:1173-1176(1992).
DR      EMBL; X62625; CAA44493.1; -.
DR      EMBL; X62626; CAA44494.1; -.
HSSP;   P02853; 2PHL.
MENDEL; 30919; Thecc:1188;30919.
DR      INTERPRO; IPR001113; -.
DR      PFAM; PF00546; Seedstore_7s; 1.
PRODOM; PD081059; -; 1.
KW      Signal.
FT      SIGNAL        1      24       POTENTIAL.
FT      CHAIN        25     525       VICILIN.
SQ      SEQUENCE   525 AA;  60798 MW;  19114CD5C248905D CRC64;

Query Match             28.9%;
Best Local Similarity   40.9%;
Matches  210;  Conservative   93;

Qy    109  ICEEEEEYNR----QRDPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKERRQQKRY  165
           :|||||||:|     |:||||||||||||||||:||||||:||:|||||:|::||:||||
Db     22  LCSGVSAYGRKQYERDPRQQYEQCQCRRCCEOSEEATEEREQCEQRCER---------EY  70

Qy    166  EEQQREDEEKYEERMKEEDNKRDPQQRYEQCQCQRRCEQQE---PRQQHQQCQLRCREQQRQH  223
           |||||||||||||:|||||||| |: |  |:|:|:|:|||||   ||  | :|::|||:|:|
Db     71  KEQQQEQEEL---------------QRQYQQCQGRCQEQQQGQRCEQQQOCRKCNWEQYKEQ  116

Qy    224  GRGGDMMNPQRGGSGRYEEGEEEEQSDNPYYFDE-RSLSTRFRTEEGHISVLENPYGRSKL  282
           ||| |:::||:| ||||     |::||||:||:| :| ||||:|||||:|| ||:|||||
Db    117  ER-GEHENYHNHKKNRSEEEEGQQRNNPYYFPRRRSFQTFRRDEEGNERKILQRAENSPP  175

Qy    283  LRALKNYRLVLLEANPNAFVLPTHLDADADILLVIGGRGALKMIHHDNRESYNLECGDVIR  342
           |:||||||:|||||||||||||||||||||  | |:|:||| |::|:|:|:||: ||| :|
Db    176  LKGINDYRLAMFEANPNTFLPHHCDAEAIYFVINGKGTITFVVTHENKESYNVQRGTVVS  235

Qy    343  IPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAAL  402
           |||||||:|:|| |:|||||||||||:|||| | |||||||||:|:|||||||||||||||
Db    236  VPAGSTVVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNKPESYIGAFSYEVLETVE  295

Qy    403  NTQTEKLRGVFGQQR------------EGVIIRASQEQIRELTRDDSESRHWHIRRGGESSRGP  454
           |||||||||||||||             |:|||||||||||:|| |||||||||:|:|||
Db    296  NTQREKLEEILEEQRGQKRQQGQGMFRKAKPEQIRAISQQATSPRH----RGGE--RLA  349

Qy    455  YNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVV  514
           ||||||||||||||||||||| ||::|||||||||||||||||||||||||||::|||||
Db    350  INLLSQSPVYSNQNGRFFEACPDEPSQFQNMDVAVSAFKLNQGAIFVVPHYNSKATFVFV  409

Qy    515  ASGEADVEMACPHLSGRHGGRGGGK--RHEEEDV----HYEQVRARLSKREAIVVLA  566
           |:||:|:||||||:||||:||||||   ||:|||     |||:|||||||||:||||
Db    410  TDGYGYAQMACPHLSRQSQGSQSQSQRDRREQEESEEEETFGEFQQVKAPLSPGDVFVAPA  469

Qy    567  GHPVFVFVSSGGNENLLLFAFGINAQNNHENFLAGR  600
           ||||||||||||||:||:||||||||||||||||||
Db    470  GHAVTFFASKDQPLNAVAFGLNAQNNQRIFLAGR  503
```

```
RESULT 7

ID   Q03865      PRELIMINARY;   PRT;   582 AA.
AC   Q03865;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE EMBRYO STORAGE PROTEIN.
GN   GLB1-L.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W64A;
RX   MEDLINE=92090707; PubMed=1752424;
RA   Belanger F.C., Kriz A.L.;
RT   "Molecular basis for allelic polymorphism of the maize Globulin-1
RT   gene.";
RL   Genetics 129:863-872(1991).
DR   EMBL; X59083; CAA41809.1; -.
DR   HSSP; P50477; 1CAU.
DR   MENDEL; 11234; Zeama; 1188;11234.
DR   INTERPRO; IPR000901; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR   PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 582 AA; 66162 MW; 680D85FECD7CB885 CRC64;

Query Match             25.4%;  Score 898;  DB 10;  Length 582;
Best Local Similarity   39.2%;  Pred. No. 3.2e-59;
Matches  197;  Conservative  91;  Mismatches 184;  Indels  30;  Gaps   9;

Qy   182 EEDNKRDPQQREYEDCRRRCEQQEPRQQHQCQLRCREQQRHGRGGDMMN-PQRGGSG-- 238
Db    25 EDDNHHHHGHKSGQCVRRCEDRRWHQRPRCLEQCREEEREKRQERSRHEADRSGEGSS   84

Qy   239 ---RYEEGEGEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLE 295
Db    85 EDEREQEKEKQKDRRPYVFDRRSFRRVVRSEQGSLRVLRPFDEVSRLLRGIRIDYRVAVLE 144

Qy   296 ANPNAFVLPTHLDADAILLVIGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRD 355
Db   145 ANPRSFVVPSHTDAHCICYVAEGEGVVTTIENGCRRSSTTIKQGHVFVAPAGAVTYLANTD 204

Qy   356 NNERLHIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLLRGVFGQ 415
Db   205 GRKKLVITKILHTISVPGFQFFPGPGGRNPESFLSSFSKSIQRAAYKTSSDRLERLFGR  264

Qy   416 --QREGVIIRASQEQIRELTRDDSE---SRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQ 470
Db   265 HCQDKGIIVRATEEQTRELRRHASEGGHGPHWPLPPPGE-SRGPYSLLDQRPSIANQHGQ 323

Qy   471 AYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSG 530
Db   324 LYEADARSFHDLAEHDVSVSFANITAGSMSAPLFNTRSFKIAYVPNGKGYAEIVCPHRQS 383

Qy   531 RHG----GRGGGKRHEEEEDVH------------------YEQVRARLSKREAIVVLAGHPVVFVSSG 576
Db   384 QGGESERERDKGRRSEEEEESSEEQEAGQGYHTIRARLSPGTAFVVPAGHPFVAVASR       443

Qy   577 NENLLFAFGINAQONHENFLAGRERNVLQQIEPQAMELAFAAPRKEVEEESFNSQDQSIF 636
Db   444 ,DSNLQIVCFEVHADRNEKVFLAGAD-NVLQKLDRVAKALSFASKAEEVDEVLGSRREKGF 502

Qy   637 FPGPRQ--HQQQSPRSTKQQQ 655
Db   503 LPGPEESGGHEEREQEEEREE 524
```

```
RESULT 8

ID   Q03678      PRELIMINARY;   PRT;   637 AA.
AC   Q03678;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   FETAL STORAGE PROTEIN.
GN   BEG1 OR GLB1.
OS   Hordeum vulgare (Barley), and Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
OX   NCBI_TaxID=4513, 4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93287988; PubMed=8510647;
RA   Heck G.R., Chamberlain A.C., Ho T.H.D.;
RT   "Barley embryo globulin 1 gene, Beg1: characterization of cDNA,
RT   chromosome mapping and regulation of expression.";
RL   Mol. Gen. Genet. 239:209-218(1993).
DR   EMBL; M64372; AAA32936.1; -.
DR   EMBL; M81719; AAA34269.1; -.
DR   HSSP; P02853; 2PHL.
DR   MENDEL; 8553; Horvu;1188;8553.
DR   INTERPRO; IPR000901; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR   PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 637 AA; 72252 MW; F323F4FF9994TC3C CRC64;

Query Match             24.4%;  Score 864.5;  DB 10;  Length 637;
Best Local Similarity   35.1%;  Pred. No. 1.2e-56;
Matches  222;  Conservative  98;  Mismatches 219;  Indels  93;  Gaps  20;

Qy    70 SKYDNQEDPQ--TECQQCQRRCRCQQESGPRQQQVCQCQRRCKEICEEEEYNRQRDPQQYE 127
Db    27 ASHDDEDDRRGGHSLQOCVQRCQRROER--PR---YSHARCVQEC-------RDDQQQH-   71

Qy   128 QCQKHCQRRETEPRIMQTCQCEBRYEKEKRKQQKRYEQQQRDEDEEKYEERMKEEDNKR  187
Db    72 ---GRHEQEEQGCRGRGWHCGEGEREEHGRGRGRHCEGEREEEHGRGRH            119

Qy   188 DPQQREYEDCRRRCEQQEPRQQHQCQLRCREQQROHGRG----GDMMNPQRGGSGRYEEG 243
Db   120 GEGEREEEERGRGHGRHGEGE-------REEERGRGRHGEGEREEEGRGRHGEG        170

Qy   244 E--EEQSDN--PYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPN 299
Db   171 ERDEEQGDSRRPYVFGRPRSFRRIIQSDHGFVRALRPFDQVSRLLRGIRIDYRVAIMEVNPR 230

Qy   300 AFVLPTHLDADAILLVIGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNER 359
Db   231 AFVVPGFTDADGVGVVAQGEGVLTVIENGKRSTVKEGDVIVAPAGSIMHLANTDGRRK   290

Qy   360 LHIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLLRGVFGQQ--- 416
Db   291 LVIAKILHTISVPGKF-QFL----SVKPLLASLSKRVLRAAFKTSDERLERLENQRQGQ   344

Qy   417 ---REGVIIRASQEQIRELTRDDSE---SRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQ 470
Db   345 EKTRSVSIVRASEQOLRELREAAEGGQGHRWPLPPFGRGDSRDTFNLLEQRPKIANRHGR 404

Qy   471 AYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVVASGEADVEMACPHLSG 530
Db   405 LYEADARSFHALANQDVRVAVANITPGSMTAPYLNTQSFKLAVVLEGEGVQIVCPHL-G  463

Qy   531 RHG---GRGGGKRHEEEED-------------------VHYEQVRARLSKREAIVVLA 566
Db   531 RHG---GRGGGKRHEEEED                                         566
```

```
Db  464  RESESEREHGKGRRREEEDDQRQQRRRGSESESESEEEEQQRYETVRARVSRGSAFVVPP  523
              |: |  |||| ||||| :| | :| :||
Qy  567  GHPVVFVSS--GNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPRKEV  624

Db  524  GHPVEISSSQQGSSNLQVVCFEINAERNERVWLAGR-NNVIGLGSPAQELTFGRPAREV  582
              ||||::|||:||| ||| || : :| :|| ||
Qy  625  EESFNSQDQSI-FPPGPRQHQQQSPRSTKQQQ  655
              :|  |||: ||||||| :|||||||||

Db  583  QEVFRAQDQDEGFVAGP---EQQSREQEQEQE  611
```

RESULT 9

ID   Q9M3X6   PRELIMINARY;   PRT;   613 AA.
AC   Q9M3X6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CONVICILIN PRECURSOR.
GN   CVC.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. BIRTE; TISSUE=SEED COTYLEDON;
RA   Casey R.;
RT   "The sequence of a pea convicilin cDNA.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ276875; CAB82855.1; -.
KW   Signal.
FT   SIGNAL   1   29   POTENTIAL.
FT   CHAIN   30   613   CONVICILIN.
SQ   SEQUENCE   613 AA;   72063 MW; 49AED99F6135DD19 CRC64;

Query Match   Best Local Similarity   23.8%;
              Score 842; DB 10;   Length 613;
Matches 212; Conservative 130; Mismatches 232; Indels 122; Gaps 15;

```
Qy    1  MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYECKRCQMQLETSGQMRRCVSQCD  60
              ||| :| | ::| :||                                 ||
Db    1  MATTIKSRFPLLLLLGIIFLASVV-----------------------------CVT---  27

Qy   61  KRFEEDIDWSKYDNQEDPQTECQCQCRCRRCRQQESGPRQQQYCCCRRCKBICEEEEEYNRQR  120
              : :|| | :                 |  |::| :: | : ::|| |  | ::
Db   28  -----YANVDEGSEPRVPAQRERGRQEGEKEEKRHGEWRPSYEKE-----EDEEGQQRER  77

Qy  121  DPQQQYEQCQHCQRRETEPRHMQTCQQRCERRYEKRKQQOKRYEEQQQREDEEKYEERM  180
              :| : :: ::          :  |   : | || |  || :|:  : :
Db   78  GRQEGEKEEKRIGEWRPS-------------------YEKQEDEEEKQKYRYQREKEDE-  121

Qy  181  KEEDNKRDPQQREYEDCRRRCEQOQEPRQQ-----------------HQCQLRCRQQROHGR  225
              | ::::           | ||||| ::| |::                |
Db  122  KYQYQREKKEQKEVQQCRERWEREEDEEQVDEEWRQSQRREDPEEARLRHRERRTKRDR  181

Qy  226  GGDMMNPQRGGSGRYEQEGEEEQSDNPYYFDERSLSTRFRTEGHISVLENFYGRSKLLRA  285
              :|   ||::: |: :||  |     :| | | |||::|: |: |||: |  |:|  |:
Db  182  ------RHQREGEEEERSSESQERRNPFLFKSNKFLTLFENENGHIHIRLLQRFDKRSDLFEN  236

Qy  286  LKNYRLVLLEANPNAFVLPTHLDADALLLVIGGRGALKMIHIDNRESYNLECGDVIRLPA  345
              |:|||| :: ||  | |: |: ||||||::|: | ||| :| || :|||| ||:|:||
Db  237  LQNYRLVEYRAKPHTIFLPQHIDADLLIVVLSGKAILTVLSPNDRNSTNLERGDTIKLPA  296

Qy  346  GTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQ  405
              |||:|| : :::| | : | : |:  | |   | :  | ||||| | | || ||| |:
Db  297  GTTSYLVNQDDEEDLRLVDLVIPVNGPGKFEAFDLA--KNKNQYLRGFSKNILEASYNTR  354

Qy  406  TEKLRGYFGGQQRE------------------GVIIRASQEQIRELTRDDSESRHWHIRRGGESSRG  453
              || : |  :|::|                    ::|| :|:|||:| :   || | |  |:|
Db  355  YETIEKVLLEEQEKDRKRRQQGEETDAIVKVSREQIELKKLKAKSSS---KKSLPSEFE  410
```

```
Qy  454  PYNLFNKRPLYSNKYGQAYEVKPE-DYRQLQDMDLSVFIANVTQGSSMMGPFFNTRSTKVV  512
              |:|| :|: |||||:| | ||: | |:||||:|| | | : | |    ||:|  |:||
Db  411  PINLRSHKPEYSNKFGKLFEITPEKKYPQLQDLDLFVSCVEINEGALMLPHYNSRAIVVL  470

Qy  513  VVASGEADVEMACPHLSGRHGGRGGGKRHEEEE--------DVHYEQVRRALSKREAIVVL  565
              :| :| :| :|     : : :|  :|||:    : ::| :|| |  |
Db  471  LVNEGKGNLELL------------GLKNEQQERDRKERNNEVQRYEARLSPGDVVIIP  517

Qy  566  AGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPRKEVE  625
              ||||| : :| ||| : |||||| ||: ||| |:|:|||| |: :|| ||||   ||:
Db  518  AGHPVAITASSNLLL--GFGINAENNERNFLSGSDDNVISQIENPVKELTFPGSVQEIN  575

Qy  626  ESFNSQDQSIFPPGPRQHQQQSPRSTKQQQPLVSTL  661
              |||||||||| ||||||| |: ||  |::
Db  576  RLIKNQKQSHFANAEPEQKEQG--SQGKRSPLSSIL  609
```

RESULT 10

ID   O22120   PRELIMINARY;   PRT;   544 AA.
AC   O22120;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ALPHA SUBUNIT OF BETA CONGLYCININ.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. WASESUZUNARI;
RA   Maruyama N., Katsube T., Wada Y., Utsumi S.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB008678; BAA23360.1; -.
DR   MENDEL; 25074; Glyma.;1188.;25074.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
SQ   SEQUENCE   544 AA;   63296 MW; O48B243929A9C9AB CRC64;

Query Match   Best Local Similarity   22.9%;
              Score 812; DB 10;   Length 544;
Matches 188; Conservative 126; Mismatches 197; Indels 70; Gaps 13;

```
Qy  107  KEICEEEEEYNRQRDPQQQYEQCQHCQRRETE--------------------PRHMQTCQQRCERRYEKE  157
              || :|| | :|| |:| : :: ::                          |||||||||:|||||||
Db    4  KEECEEGEIPRPRPRPQHPEREPQQPGEKEDEDEDEQPRPLPFFPRPQPQEEEHEQREEQE  63

Qy  158  -KRKQQOKRYEEQQQREDEEEKYEERMKEEDNKRDPQQOREYEDCRRRCEQOQEPRQQHCQLRC  216
              : |  || |: |:: ::   : :| : | :           : | ::  |   |:
Db   64  WPRKEEKRGEKGSEEEDE--------DEDEQEDERQFPFP---RPPHQKEERKQEEDE--D  111

Qy  217  REQQRQHGRGGGDMMNPQRGGSGRYEQEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENF  276
              ||| : :                   : :|  |||:   ::|  |:: :|:  |
Db  112  EEQQRESEESED---------------SELRRHKNKNPFLFGGSNREETLFKNQYGRIRVLQRF  159

Qy  277  YGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHIDNRESYNLE  336
              :|||  |  | ||| :: ||  |  | |: ||||:|:||:| :|||| :|:|| :|||
Db  160  NQRSPQLQNLRDYRILEFNSKPNTLLPNHHADADYLIVLNGTAILSLVNNDRDDSYRLQ  219

Qy  337  CGDVIRIPAGTTEYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKE  396
              |||::|| |||| |: : :|||| | : | : |:  | |   | |  | ||||| | |:
Db  220  SGDALRVPSGTTYYVVNPDNNENLRLITLAIPVNKPGRESEFLLSSTEAQQSYLQGFSRN  279

Qy  397  ILEAALNTQTEKLRGVF------------------GQQR--EGVIIRASQEQIRELTRDDSESRHWHIR  445
              |||| : :| || |   :                   |:|| |:|| :|:||| |  :|| | |
Db  280  ILEASYDIKFEEINKVLFSREEGCQOGEQORLQESSVIVEISKEQIRALSKRAKSSS----R  335
```

```
Qy   446  RGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFN  505
               |  ||||  : || :| ||||| | | |  :||||   :||::||:|  || :| || |
Db   336  KTISSEDKPFNLRSRDPIYSNKLGKFFEITTPEKNPQLRDLDIFLSIVDMNEGALLLPHFN  395

Qy   506  TRSTKVVVVASGEADVEMACPHLSGRHGGRGGGK-----RHEEEEDVHYEQVRARLSKRE  560
                               :||||||||||  :||     |::|: : |  | ||:|: |
Db   396  SKAIVLIVINEGDANIELV-------GLKEQQOQOQEQPLEVRKYRAELSEQD         443

Qy   561  AIVVLAGHPVVFVSSGNENLLLFAFGINAQNHHENFLAGRERNVLQQIEPQAMELAFAAP  620
          :| :||| |||  :||::||   |:|||| :: ::| || | :|| | | |: | |||| |
Db   444  IFVIPAGIPVVV--NATSNLNFFAIGINAENNQRNFLAGSQDNVISQIPSQVQELAFPGS  501

Qy   621  RKEVEESFNSQDQSIFFPGPROHQQOSPRSTKQQOPLVSIL                     661
               :| |:|:|| ||||||| |    :|: | ||| ||:|||
Db   502  AQAVEKLLKNQRESYFVDAQPKKKEEGNKGRK--GPLSSIL                     540


RESULT 11
Q9SP11

ID   Q9SP11       PRELIMINARY;      PRT;   489 AA.
AC   Q9SP11;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   SUCROSE BINDING PROTEIN HOMOLOG S-64.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pedra J.H.F., Delu-Filho N., Pirovani C.P., Contim L.S., Dewey R.E.,
RA   Otoni W.C., Fontes E.P.B.;
RT   "Antisense and sense expression of a sucrose binding protein homologue
RT   gene from soybean in transgenic tobacco affects plant growth and
RT   carbohydrate partitioning in leaves.";
RL   Plant Sci. 0:0-0(1999).
DR   EMBL; AF191299; AAF05723.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE   489 AA;  55834 MW;  9BBC0D45EDECECD2 CRC64;


Query Match             22.7%;  Score 802.5;  DB 10;  Length 489;
  Best Local Similarity   35.8%;              Pred. No. 3.6e-52;
  Matches  175;  Conservative   98;  Mismatches  171;  Indels   45;  Gaps   11;

Qy   179  RMKEEDNKRDP---------QQREYED------CRRRCEQOEPROQHQCOLRCREQOQROH  223
            | ||: || |         :|:|      ::|| || |: :  |:|   ||:|::|:: |
Db    27  KLKETEVEEDPELVTCKHQCOOQORQYTESDKRTCLQOCDSMKQEREKQVEEETREKEEEH   86

Qy   224  GRGGDMMNPQRGGSGRYEEGEGEEEQSDNPYYFDE-RSLSTRFRTEEGHIISVLENFYGRSKL  282
                     :|::|: |: ||: ::: ||| | || |: || || |:||  :||||||||
Db    87  ----------QEQHEEEQDQNPYVFEEDKDFSTRVETEGGSIRVLKKFTEKSKL          130

Qy   283  LRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMHHDNRESYNLECGDVIR     342
           | :::| || |:|| |  ||:|:||| | ::|:| |:|| | ||| |||:| | ||:|
Db   131  LOGNENFRLAILEARAHTFVSPRHFDSEVVLLNIKGRAVLGLVRESETEKITLEPGDMIH    190

Qy   343  IPAGTTFYLINRDNNERLHIAKFLQTISTPGGYKEFFPAGGONPEPYLSTFSKEILEAAL    402
          |||| ::|:|:||:||:|  | :|:|:|:|| |:|||| || :||||||:||  |: ||
Db   191  IPAGTPLYIVNRDENEKLLLAMLHIPVSTPGKFEEFFGPGGRDPESVLSAFSWNVLQAAL    250

Qy   403  NTQTEKLRGVFCGQQREGVIIRASQEQIRELTRDDSESRHWHIRRGGESSSRGPYNLFNKRP  462
            | :|| |:  :| |  |::| |:|:||  |: |::| | : | |::: |: |||| ||||
Db   251  QTPKGKLERLENQQNEGSIFKISRERVRALA-PTKKSSWWPF--GGE-SKAQFNIFSKRP    306

Qy   463  LYSNKYGQAYEVKPEDYRQ-LQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADV   521
          ::|| ||:   :| ||::  |: :|| |:: |:|| || ::: : : |:| :|:      |
Db   307  TFSNGYGRLTEVGPDDEKSWLQRLNLMLTFTNITQRSMSTIIHYNSHATKIALVMDGRGHL  366
```

```
Qy   522  EMACPHLSGRHGGRGGGKRIHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLL  581
          ::| |: |||:|:||   |::| :: :  |  :: |:  | |:||:||||| ||:||:||||
Db   367  QISCPHMSSRSDSK-----HDKSSPSYHRISADLKPGMWFVVPGCHPEVTIASNKENLL    420

Qy   582  LFAFGINAQNHHENFLAGRERNVLQQIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPR   641
          : ||::|::|:: |||| ||| |:  || ||| |||  |::: ::||:  :: :|||| |
Db   421  IICFEVNVRDNKKFFFAGKD-NIVSSLDNVAKELAFNYPSEMVNGVF-ERKESLFFFPEL   478

Qy   642  QHQQOSPRS                                                     650
           : || ||
Db   479  PSEERGRRA                                                     487


RESULT 12
Q41674

ID   Q41674       PRELIMINARY;      PRT;   545 AA.
AC   Q41674;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CONVICILIN PRECURSOR.
OS   Vicia narbonensis.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Vicia.
OX   NCBI_TaxID=3912;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COTYLEDON;
RA   Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z71986; CAA96513.1; -.
DR   HSSP; P50477; 1CAU.
DR   MENDEL; 12432; Vicna;1188;12432.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
KW   Signal; Seed storage protein.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    545       7S GLOBULIN, CONVICILIN.
SQ   SEQUENCE   545 AA;  62810 MW;  459A876F92F5A87E CRC64;


Query Match             22.2%;  Score 787;  DB 10;  Length 545;
  Best Local Similarity   33.5%;              Pred. No. 6.1e-51;
  Matches  191;  Conservative  102;  Mismatches  208;  Indels   70;  Gaps   14;

Qy   109  ICEEEEEYNRQRDPQOQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKERKQQKRYEE-   167
           | :::::| ||:|   : :  :: || :|     | |  : :|: ::|   |: ::
Db    23  LCVTYANYDEGTEPRVPGQRERGRQEGEKEEKRH----------GEWRPSHEKEAQPGRRERW   75

Qy   168  QQREDEEKYEERMKEEDNKRDPQQREYEDCRRRCEQOEPRQQHQCOLRCREQOROHGRGG   227
          : :|:||:: ||  :|: || | :|  |  || |: :|   :: |:|:| :: :  ||
Db    76  ETSEEEERVDEEWRGSQRIIEDPEERARE--RYRAEEERRRQWE-------------    117

Qy   228  DMMNPQRGGSGRYEEGEEEQSD--NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRA  285
                :::|| :|: ||: :: |  || :| || || ||||| |:|||||||||||| | |
Db   118  ------GEEKEGSSKSQERRNPFLFKSNKFLTLFENENGHIRRLQRFDKRSDLFEN     167

Qy   286  LKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVIRIPA  345
          |:|||| :   : :  ||| ::|||: : | | ::|| | ||| | | ||| ||:| ||
Db   168  LQNYRLVEYRAKPHTIFLPQHIDADLITVLSPNDRNSYNLERGDTIKLPA           227

Qy   346  GTTFYLINRDNNERLHIAKFLQTISTPGGYKEFFPAGGONPEPYLSTFSKEILEAALNTQ  405
          || :|:|:|| :|:|  | :| ||| | | ||| |:|| :|  ||:| |||: ||:|||:
Db   228  GTISYLLNQDDDEDLRVDLSISVNRPGKVESFGLSGSKN--QYLRGFSKNILEASLNTK   285

Qy   406  TEKLRGV--------FGQQRE-----GVIIRASQEQIRELTRDDSESRHWHIRRGGESS   451
           ||:      ||:|      |:::|:|:|  |: |::| | : | |::: |
Db   286  YETTEKVLLEEPQOSIGQKRRSQRQETNALVKYSREQVEELKIRLAKSSS----KKGVSSE  341

Qy   452  RGPYNLFNKRPLYSNKYGQAYEVKPE-DYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTK  510
```

(continuation of previous alignment)

```
Db  342 FEPFNLRSQNPKYSNKPGKLFEITPERKYYPQLQDLDIFVSSVEINEGGLMLPHYNSRAIV 401
Qy  511 VVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPV  570
Db  402 ILLVNEGKGNLE-----LVGLKNEQEQREDEEQQV--QRYEARLSPGDVVIIPAGHPV     454
Qy  571 VFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPKREVEESFNS   630
Db  455 AVSASSNLNLL--GFGINAENNQRNFLTGSDDNVISQIENPVKELTPGSAQEVNRLLKN    512
Qy  631 QDQSIFFPGPRQHQQQPSPRSTKQQQPLVSIL 661
Db  513 QEHSHF--ANAPEQKGEESQQKRSPISSIL 541
```

RESULT 13

```
ID   O49927
AC   O49927;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   P54 PROTEIN.
GN   P54.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y11207; CAA72090.1; -.
DR   HSSP; P50477; 1CAU.
DR   MENDEL; 28319; Pissa;3005.28319.
DR   INTERPRO; IPR000005; -.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
DR   PROSITE; PS00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
SQ   SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

Query Match   21.7%;  Score 769.5;  DB 10;  Length 483;
Best Local Similarity  31.4%;
Matches 162;  Conservative 114;  Mismatches 169;  Indels 71;  Gaps 12;
```

```
Qy  135 RRETEPRHMQTCQQRCERRYEKERKKQQKRYEEQQREDEEKYEERMKEEDNKRDPQQREY 194
Db   28 RKEKDP-ELTTCKDQCD------------MQRQVDEEDK----------------------  53
Qy  195 EDCRRRCEQQEPRQQHQCQLRCREQQRQHGRGGDMMNPQRGGSGRYEEGEEGEEQSDNPYYF 254
Db   54 RICMERQDYIKKQE-----------------RQKHKEH------------------------  93
Qy  255 DERSLSTRFRTEEGHHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILL  314
Db   94 EDNDFETKIDTKDGRVLILNKFNEKSLLKNIENYGLAVLEIKANAFLSPHHYDSEAILF    153
Qy  315 VIGGRGALKMIHHDNRESYNLECGDVIRIPAGTFFYLINRDNNERLHIAKF---LQTIST   371
Db  154 NIKGRGIIGLVAEDRTRERFNLEEGDIMRVPAGTPMYLVNRDENEKLYIAAFHMPPSSGSA  213
Qy  372 PGQYKEFFPAGGQNPEPYLSFSKEILEAALNTQTEKLRGYFGQQOREGVIIRASQEQIRE   431
Db  214 PVNLEPFFESAGRKPESVLNTFSSKVLQAALKSSKGELETVLDEQKGRIFKIEKEDVRG    273
Qy  432 LTRDDSESRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQ-LQDMDLSVF   490
Db  274 LAPKKS---LWPF---GGPFKSPENIFSNNPAFSNKFGSLFEVGPSQOEBKSGLEGLNLMLT 327
```

RESULT 14

```
Qy  491 IANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEDVHYE  550
Db  328 LANITKGGSMSTHYNTNANKIALVIDGEGELEMACPHMPS-----SSSNSRQKKSSISYH  382
Qy  551 QVRARLSKREAIVVLAGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEP  610
Db  383 NINAKLRPGVMFVVVPAGHPFVNIASKKNLIVCFEVNAQRNKKLALAGK-KNIVSALDK   441
Qy  611 QAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQQQ 646
Db  442 AAKEVAFDIAAEKVDEVFERKEE-FFFPYDNEERKE 476
```

```
ID   Q40873
AC   Q40873;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE STORAGE PROTEIN.
OS   Picea glauca (White spruce).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX   NCBI_TaxID=3330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PG11B; TISSUE=SOMATIC EMBRYO;
RX   MEDLINE=93004485; PubMed=1391775;
RA   Newton C.H., Flinn B.S., Sutton B.C.;
RT   "Vicillin-like seed storage proteins in the gymnosperm interior spruce
RT   (Picea glauca/engelmanii).";
RL   Plant Mol. Biol. 20:315-322(1992).
DR   EMBL; X63191; CAA44873.1; -.
DR   HSSP; P02853; 2PHL.
DR   MENDEL; 12115; Picgl;1188.12115.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 448 AA; 50200 MW; 467AE5BA12E6D157 CRC64;

Query Match   21.2%;  Score 752;  DB 10;  Length 448;
Best Local Similarity  37.8%;
Matches 162;  Conservative 85;  Mismatches 158;  Indels 24;  Gaps 9;
```

```
Qy  232 PQRGGSGRYEEGEEGEEQSDNPYYFDERSLSTRFRTEEGHHISVLENFYGRSKLLRALKNYRL 291
Db   34 PEYLGRGRGRR-EBEREENPYVHSDSRFTRASSEAGEIRALNFGEVSELLEGIRKFRV       92
Qy  292 VLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVIRIPAGTFFYL    351
Db   93 TCIEMKPNTVMLPHYIDATWILVYTRGGYIAYVHQNELVKRKLEEGDVFGPSGHTFYL     152
Qy  352 INRDNNERLHIAKFLQTIST-PGQYKEFFPAGGQNPEPYLSFSKEILEAALNTQTEKLR    410
Db  153 VNDDHNTLRIASLVRPVSTVRGEYQPFYVAGGRNPQTYVSAFSDDVLEAAFNTNVVQOLE   212
Qy  411 GVFGQQOREGVIIRASQEQIRELTRDDSESRHWHIRRG---GESSR----GPYNLFNKRPL  463
Db  213 RIFGGHKSGVIIHANEEQIREMMR--------------KRGFSAGSMSAPEHPKPFNLRNQKPD 262
Qy  464 YSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEM   523
Db  263 FENENGRFTIAGPKNYPFLDALDVSVGLADLNPGSMTAPSLNSKSTSIGIVINGEGRIEM   322
Qy  524 ACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLLF   583
Db  323 ACPHLGQIG--WSSPRERGDQQITYQRVWAKLRTGSVIVPAGHPITEIASTNSRLQIL     379
Qy  584 AFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPR-KEVEESFNSQDQSIFFPGPRQ   642
```

```
Db    380 WFDLNTRGNERQFLAGK-NNVLNTLEREIRQLSFNVPRGEEIEEVLQAQKDQVILRGPQR 438

Qy    643 HQQQSPRST 651
              ||:|
Db    439 RSRDEARSS 447
```

```
RESULT 15
ID    Q9LUJ7         PRELIMINARY;    PRT;    486 AA.
AC    Q9LUJ7;
DT    01-OCT-2000 (TrEMBLrel. 15, Created)
DT    01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT    01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE    GENOMIC DNA, CHROMOSOME 3, P1 CLONE: MWI23.
OS    Arabidopsis thaliana (Mouse-ear cress).
OC    Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC    Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC    Brassicales; Brassicaceae; Arabidopsis.
OX    NCBI_TaxID=3702;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=COLUMBIA;
RA    Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL    Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=COLUMBIA;
RA    Nakamura Y.;
RT    "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT    features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT    clones.";
RL    DNA Res. 7:131-135(2000).
DR    EMBL; AB022223; BAB01239.1; -.
SQ    SEQUENCE  486 AA;  55063 MW;  325ECF68D9A6345B CRC64;

Query Match     21.1%;  Score 749;  DB 10;  Length 486;
Best Local Similarity   35.9%;  Pred. No. 3.7e-48;
Matches  174;  Conservative   92;  Mismatches 179;  Indels   40;  Gaps  12;
```

```
Qy    198 RRRCEQQEPRQQHQCQLRCREQQROHGRGGDMMNPQR---GGSGRYEEGEEEQSDNPYYF 254
              :-|||  |:|||:||||  |:-              ||-:-::: ||::||||:||| ||
Db     26 RVEIDEFEPPQQ----------------------GEQEGPRRRPGGGSS--EGWEEESTNHPYHF 66

Qy    255 DERSLSTRFRTEEGHISVLENFYGRS-KLLRALKNYRLVLLEANPNAFVLPTHLDADAIL 313
              :-|||  |  ::||||:|  | ||| ::||| |:|:|||  |:| | |||||||:||||:
Db     67 RKRSFSDWFQSKEGFVRVLPKFTKHAPALFRGIENYRFSLVEMEPTFFVPHHLDADAVF 126

Qy    314 LVIGGRGALKMIHIDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKFLQTISTPG 373
              :|:|  |:|  |  || ||:| ||||:|||||||||||||:|:|||||||||||||||||
Db    127 IVLQGKGVIEFVIDKTKESFHITKGDVVRIPSGVTNFITNTNQTVPLRLAQITVPVNNPG 186

Qy    374 QYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQRE---GVIIRASQEQIR 430
              ||:|||||||:|||||:|  ||:||||:|:||:|:|||||||||   :||||||:|||||
Db    187 NYKDYFPAASQFQQSYENGFTKEVLSTSFNVPEELLGRLVTRSKEIGQGIIRRISPDQIK 246

Qy    431 ELTRDDSESRHWH-IRRGGESSRG-----PYNLFNKRPLYSNKYGQAYEVKPEDYRQLQ 483
              ||  | ::|:|||  ||||:|||      ||||||||||:||||||||:||:|:|||||
Db    247 ELAEHATSPSNKHKAKEEEKDLRTLWTPENLFAIDPIYSNDFGHFHEAHPKNYNQLQ 306

Qy    484 DMDLSVFIANVTQGSMMGPFFNTRSTKVVVASGEADVEMACPHLSGRHGGRGGGKRHEE 543
              |:-||  |  :||  |: ||||| ||:|:|:|||:|:|||:||:||:||:||||:|||||
Db    307 DLHIAAANAMTQGSLFLPHFNSKTTFVTFVENGCARFEMATPYKFRQRQQQOWPGQGQEE 366

Qy    544 EEDV--HYEQVRARLSKREAIVVLAGHPVFVFVSSGNENLLFAFGINAQNNHENFLAGRE 601
              |:|:  :|  |:|  ||| |:|:|  ||||:|||:|:|||:|||||||||:||||:|:||
Db    367 EEDMSENVHKVSRVCKGEVFIVPAGHPFTILSQ-DQQFIAVGFGIYATNSKRTFLAG-E 424

Qy    602 RNVLQQIEPQAMELAFAAAPRKEVEEESFNSQDQSIFFPGPROHQQQSPRSTKQQOPLVSIL 661
              ||:||  ||: | |||  | |  |:|||:||||||:|||||:||||||:|:||:|||||:||
Db    425 ENLLSNLNPAATRVTFGVGSKVAEKLFTSQNYSYFAPTSRS-QQQIPE-KHKPSFQSIL 481
```

```
Qy    662 DFVGF 666
              |:| ||
Db    482 DFAGF 486
```